PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 : C12N 15/12, C07K 14/435 | A1

(11) International Publication Number: **WO 00/15789**

(43) International Publication Date: 23 March 2000 (23.03.00)

(21) International Application Number: PCT/US99/20463

(22) International Filing Date: 7 September 1999 (07.09.99)

(30) Priority Data:
09/149,973 9 September 1998 (09.09.98) US

(71) Applicant: UNITED STATES SURGICAL CORPORATION [US/US]; 150 Glover Avenue, Norwalk, CT 06856 (US).

(72) Inventors: PAOLELLA, David, N.; 166 Route 81 # A, Killingworth, CT 06419 (US). GRUSKIN, Elliott, A.; 23 Beach Tree Ridge, Killingworth, CT 06419 (US). BUECHTER, Douglas, D.; 333 Orchard Street, Wallingford, CT 06492 (US).

(74) Agent: FARBER, Mark; United States Surgical Corporation, 150 Glover Avenue, Norwalk, CT 06856 (US).

(81) Designated States: AL, AM, AT, AU, AZ, BB, BG, BR, BY, CA, CH, CN, CZ, DE, DK, EE, ES, FI, GB, GE, HU, IL, IS, JP, KE, KG, KP, KR, KZ, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TM, TR, TT, UA, UG, UZ, VN, European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: RECOMBINANT BIOADHESIVE PROTEIN ANALOGS COMPRISING HYDROXYPROLINE

(57) Abstract

A method for the production of bioadhesive precursor protein analogs having incorporated therein hydroxyproline, wherein vectors comprising DNA sequences encoding the bioadhesive protein precursor analog are constructed and used to transform a microbial host, preferably *E. coli*. During growth, the host is contacted with a hypertonic media comprising hydroxyproline, resulting in assimilation and incorporation of hydroxyproline into the bioadhesive precursor protein during translation. Incorporation of hydroxyproline into the bioadhesive precursor proteins during protein production is advantageous in that post–translational modification of the proteins are not required.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

RECOMBINANT BIOADHESIVE PROTEIN ANALOGS COMPRISING HYDROXYPROLINE

Background

1. Field

Engineered polypeptides having bioadhesive properties, in particular, a method for the production of bioadhesive polypeptides wherein hydroxyproline and/or other proline analogs is incorporated into the polypeptide at the translational level, thereby obviating the need for post-translational proline hydroxylation.

2. Description of the Related Art

Marine mussels and other sessile invertebrates such as platyhelminths, annelids, and tunicates secrete adhesive substances in order to affix themselves to underwater objects. Mussels of the genus *Mytilus*, for example the species *Mytilus edulis* and *Mytilus californianus*, secrete an adhesive precursor substance from their foot that upon cure forms a permanent attachment to the substrate. A major component of the adhesive deposited by *M. edulis* has been identified as a hydroxylated protein of about 130,000 daltons (Waite, J. H., J. Biol. Chem., Vol. 258, pp. 2911-2915 (1983)). U.S. Pat. No. 4,496,397 to Waite further discloses that the protein from *M. edulis* comprises a large number (75-80) of tandem repeats of a decapeptide having the following sequence:

Ala-Lys-Pro/Hyp-Ser/Thr-Tyr/Dopa-Pro/Hyp-Pro/Hyp-Ser/Thr-Tyr/Dopa-Lys

According to U.S. Pat. No. 5,574,134 to Waite, proteins derived from other mussel species comprise repeating tripeptides, octapeptides, and nonapeptides having the following sequences:

Gly-Dopa-Lys

Thr/Ala-Gly-Dopa/Tyr-Y^1 -Y^2 -Gly-Dopa/Tyr-Lys

Gln-Thr/Ala-Gly-Dopa/Tyr-Y^1 -Y^2 -Gly-Dopa/Tyr-Lys

wherein Y^1 is a serine, valine, aspartic acid, or leucine residue and Y^2 is an alanine, proline, hydroxyproline, or leucine residue. In U.S. Patent No. 5,410,023, Burzio discloses that the protein from the mussel *Aulecomya ater* comprises a repeating heptapeptide having the sequence:

Ala-Gly-Tyr/Dopa-Gly-Gly-Val/Leu/Iso-Lys/LysOH

Precursor proteins comprising the above sequences have a large number of basic and hydroxylated amino acids, including lysine, serine, threonine, and others. The relatively large numbers of basic and hydroxylated amino acid presumably provide both hydrogen and ionic bonding with the substrate surface. At least a portion of the hydroxylated amino acids include 3,4-dihydroxyphenylalanine (also know as dopamine, abbreviated as "Dopa") and hydroxyproline (abbreviated as "Hyp"). Dopamine and hydroxyproline are unique among amino acids, because they are incorporated into the protein as tyrosine and proline during protein synthesis ("translation"), and then later hydroxylated. This process is often referred to as "post-translational modification", or in this instance, "post-translation hydroxylation". Post-translational hydroxylation, particularly of the tyrosine residues, is believed to be important in defining the adhesive properties of the protein. The hydroxyl groups likely

chelate with metals or other ions present on the substrates on which mussels anchor. Oxidation of the dihydroxy groups of the tyrosine yield ortho-phenols, which can then cross-link and form or strengthen the adhesive bond. See, Waite, J. H., In Mollusca, Vol. 1, pp.467-504 (1983); Pizzi, A., et al., Ind. Eng. Chem. Prod. Res. Dev., Vol. 21, pp. 309-369 (1982) and Wake, W.C., "Adhesion and the Formulations of Adhesives", Applied Science Publish, Ltd. Barking, England (1982).

Mussel adhesive proteins have been the subject of intensive study because they are anticipated to provide a natural adhesive for use in wet environments. Wet environments are found in various areas of clinical medicine such as opthamology, cardiology, otology, neurology, orthopedics, orthodontics, and other related fields. Ideal adhesives for these environments are biocompatible, curable in wet environments, have adhesive strengths comparable to sutures or staples, biodegradable, hemostatic, and do not interfere with normal healing.

Adhesives and sealants in current therapeutic use include both proteinaceous and synthetic formulations. One proteinaceous formulation is a mixture of the human proteins fibrinogen, thrombin, and Factor XIII, known as fibrin sealant, or "fibrin glue". Although fibrin glue is biocompatible, hemostatic, and does not interfere with normal healing, it suffers from poor mechanical strength. Synthetic cynaoacrylate adhesives have superior mechanical strength compared to fibrin glue but are not completely biocompatible, causing inflammation in non-surface applications. Other bioadhesives include prolamine gels, which are based on zein proteins in plants, and which are useful for pancreatic duct obstruction and venous angiomas. Gelatin resorcinol formol (GRF) glues are useful for cardiovascular, hepatic, and

renal applications. However, none of the aforementioned products is suitable for a wide range of applications, and with the exception of fibrin glue, none has entered routine use.

Mussel adhesive proteins have also failed so far to live up to their initial promise. This is for a variety of reasons, including difficulties in obtaining large quantities of the protein precursor, and in obtaining efficient cure of the protein precursor. For example, isolation of the uncured adhesive precursor from mussels for commercial use is not practical. Extraction, isolation, and purification is labor-intensive, and calculations indicate that extraction of even about 1 kilogram of the adhesive precursor would require about 5 to 10 million mussels.

Accordingly, a number of groups have investigated the utility of semi-synthetic, synthetic, and recombinant routes to obtain the larger protein quantities required for commercial applications. U.S. Pat. Nos. 4,585,585 and 4,808,702 to Waite describe a semi-synthetic procedure for preparing a bioadhesive polymer by chemically linking decapeptide units produced by the enzymatic digestion of isolated mussel adhesive protein. Thus, a bioadhesive protein is first isolated from phenol glands of mussels of the genus *Mytilus*. The isolated protein, having a molecular weight of 120,000 to 140,000 daltons, is treated with collagenase, and then with trypsin, to yield, after gel filtration and dialysis, the decapeptide disclosed by Waite in U.S. Pat. No. 4,496,397. These decapeptides are then polymerized, using bifunctional linking groups such as glutaraldehyde, oligopeptides, and amino acids, producing molecules having up to about 1,000 decapeptide units. This procedure still requires the isolation of bioadhesive protein from mussels, which, as previously indicated, is impractical on a commercial scale. Moreover, in addition to the laborious protein

purification procedure, the steps of enzymatic digestion, isolation of the decapeptide fragments, and chemical reassemblage of the fragments into a bioadhesive polymer have been added.

U.S. Pat. No. 4,908,404 to Benedict et al. discloses a synthetic approach to a mussel adhesive protein analog. A graft co-polymer comprising a poly(lysine), poly(allylamine), chitosan, polydextran, or similar backbone, is grafted with various peptides or the decapeptide disclosed by Waite '397, supra. The protected decapeptides are obtained by solid-phase synthesis. The disclosed synthetic steps are complex, require expensive starting materials, and are not amenable to commercial scale-up.

Recombinant methods for obtaining large quantities of mussel adhesive proteins have also been developed, as disclosed in U.S. Pat. Nos. 5,049,504, 5,149,657, 5,202,236, 5,202,256, and 5,242,808 to Maugh et al. These patents disclose the application of the techniques of recombinant DNA technology to the production of bioadhesive precursor proteins of the type produced by *M. edulis*. In these approaches, the genes coding for the mussel adhesive precursor protein is cloned into a simple organism such as a bacteria, and the bacteria's own cellular machinery is used to make ("express") the protein in large quantities. The bacteria will usually then secrete the protein, which can be isolated and purified. U.S. Pat. No. 5,149,657, for example, discloses construction of a DNA encoding a bioadhesive precursor protein analog. The analog comprises three, ten, fifteen or twenty repeating decapeptides of the sequence

Ala-Lys-Pro-Ser-Tyr-Pro-Pro-Thr-Tyr-Lys

The DNA is inserted into a plasmid (the "expression vector"), and the plasmid is then inserted into a host (referred to as "transformation"). During growth, the host is induced to produce the precursor protein analog, which is then isolated. At least a portion of the tyrosine residues of the isolated protein are then hydroxylated by mushroom tyrosinase to form dopamine, and the hydroxylated protein is cured to produce the desired physical properties in the adhesive.

In U.S. Pat. No. 5,202,236, genes for native bioadhesive precursor proteins of the mussel *M. edulis*, together with constructed expression vectors capable of expressing these precursor proteins in microorganisms, are cloned in microbial hosts. After isolation from the microbial broth, at least a portion of the tyrosine residues of the bioadhesive precursor proteins are then enzymatically hydroxylated by mushroom tyrosinase, mimicking the process which occurs in the mussel. U.S. Patent No. 5,242,808 discloses an alternative scheme for producing analogs of mussel bioadhesive precursor proteins, first comprising synthesis of double-stranded DNA (dsDNA) encoding proteins having from about 50 to about 1500 amino acid residues. These proteins also have from about 20% to 40% proline residues; from about 10% to 40% lysine residues; from about 10% to a 40% tyrosine residues; and from 0 to 40% other amino acid residues. The dsDNA is inserted into a vector, and the vector is used to transform a host which is then used to produce the protein. Finally, the tyrosine residues of the produced protein are at least partially hydroxylated to dopamine by mushroom or *Streptomyces antibioticus* tyrosinase prior to cure.

A significant drawback of the recombinant methods disclosed in the above-mentioned references is that none of the produced proteins comprise hydroxyproline, an integral

component of naturally-occurring bioadhesive proteins. It is well known that certain cells incorporate a complex set of post-translational enzyme systems which impart unique characteristics to protein products of the systems, in particular, post-translational hydroxylation of tyrosine to dopa, and proline to hydroxyproline. When a gene encoding a protein normally produced by these cells is transferred into a bacterial or yeast cell (such as *E. coli* or *S. cervisiae*), the protein may not be subjected to such post transitional modification and the protein may therefore not function as originally intended. In the present instance, the presence of hydroxyproline likely contributes to the tertiary structure of the natural protein, as well as to the formation of hydrogen and/or ionic bonds during adhesion.

Although proline in the produced proteins could conceivably be hydroxylated by chemical or enzymatic means to hydroxyproline, such transformation would add at least one additional step the preparation of biodhesive precursor proteins. The inventors hereof are furthermore presently unaware of any references describing such post-translational modification of bioadhesive proteins. Accordingly, there remains a need in the art for methods of producing bioadhesive precursor proteins having incorporated therein hydroxyproline and/or other proline analogs efficiently, and using techniques amenable to commercial scale.

Summary

The above-discussed and other drawbacks and deficiencies of the prior art are overcome or alleviated by a method comprising the production, using recombinant techniques, of bioadhesive precursor protein analogs having incorporated therein hydroxyproline. In accordance with the method, vectors comprising DNA sequences encoding the bioadhesive protein precursor analog are constructed and used to transform a microbial host, preferably *E. coli*. During growth, the host is contacted with a hypertonic growth media comprising hydroxyproline, resulting in assimilation and incorporation of hydroxyproline into the bioadhesive precursor protein. Incorporation of hydroxyproline into the bioadhesive precursor proteins during protein production is advantageous in that post-translational hydroxylation of proline is no longer required. The resulting precursor protein is then isolated, purified if necessary, and the tyrosine residues hydroxylated by chemical or enzymatic means prior to cure.

The bioadhesive precursor protein analog produced by this method comprises from about 1% to about 50%, from about 3% to about 30%, or from about 5% to about 20% hydroxyproline based on the total number of amino acid residues in the protein. Other proline analogs or mixtures may also be incorporated. Preferably, the bioadhesive precursor protein analog comprises from three to about one hundred, from three to about fifty, or from three to about twenty-five repeating decapeptide units of the sequence

Ala-Lys-Pro/Hyp-Ser-Tyr-Pro/Hyp-Pro/Hyp-Thr-Tyr-Lys

wherein, within the protein, at least one of the above residues designated Pro/Hyp is hydroxyproline, other protein analog, or mixtures thereof. Alternatively, from about 1% to

about 100%, from about 5% to about 75%, from about 15% to about 50%, or from about 5% to about 15% of the Pro/Hyp residues within the protein comprising the above decapeptide is hydroxyproline, other proline analog, or a mixture thereof. Hydroxyproline analogs of proline such as *trans*-4-hydroxy-L-proline, *cis*-4-hydroxy-L-proline, *trans*-3-hydroxy-L-proline, *cis*-3-hydroxy-L-proline, and 3,4-dihydroxy-L-proline are of particular interest. However, other proline analogs and mixtures thereof which are imported into the host cell by the same mechanism may also be used, including but not being limited to *cis*-4-fluoro-L-proline, 3-fluoro-L-proline, 3,4-epoxy-L-proline, and the like.

Other embodiments comprise vectors having the DNA sequences encoding the bioadhesive protein precursor analogs, *E. coli* hosts transformed with said vectors, the bioadhesive precursor protein analogs produced thereby, and adhesives formed from said bioadhesive precursor protein analogs.

The above-discussed and other features and advantages will be appreciated and understood by those skilled in the art from the following detailed description and drawings.

Brief Description of the Drawings

Referring now to the drawings wherein like elements are numbered alike in the several FIGURES:

FIG. 1 is a dsDNA sequence encoding a bioadhesive precursor protein decapeptide having the sequence Ala-Lys-Pro/Hyp-Ser-Tyr-Pro/Hyp-Pro/Hyp-Thr-Tyr-Lys.

FIG. 2 is a flow chart showing the assembly of a repetitive DNA sequence encoding a bioadhesive precursor protein analog comprising decapeptides having the sequence Ala-Lys-Pro/Hyp-Ser-Tyr-Pro/Hyp-Pro/Hyp-Thr-Tyr-Lys.

FIG. 3 is a diagram of *E. coli* plasmid pLSM-6 containing a DNA sequence encoding 25 repeats of a bioadhesive precursor protein decapeptide having the sequence Ala-Lys-Pro/Hyp-Ser-Tyr-Pro/Hyp-Pro/Hyp-Thr-Tyr-Lys.

FIG. 4 is a diagram of *E. coli* plasmid pLSM-10 containing a DNA sequence encoding 25 repeats of a bioadhesive precursor protein decapeptide having the sequence NH_2-Ala-Lys-Pro/Hyp-Ser-Tyr-Pro/Hyp-Pro/Hyp-Thr-Tyr-Lys.

FIG. 5 is a flow chart showing the assembly of a repetitive DNA sequence encoding a bioadhesive precursor protein analog comprising decapeptides having the sequence Ala-Lys-Pro/Hyp-Ser-Tyr-Pro/Hyp-Pro/Hyp-Thr-Tyr-Lys and linked by hexapeptides having the sequence Lys-Tyr-Pro/Hyp-Ser-Ala-Lys.

FIG. 6 shows by SDS-PAGE analysis expression of protein precursor analogs comprising *trans*-4-L-hydroxyproline in pLSM-10 in *E. coli*.

FIG. 7 shows by SDS-PAGE analysis expression of protein precursor analogs comprising *trans*-4-L-hydroxyproline (Hyp) and *cis*-4-hydroxy-L-proline (cis-4-Hyp) in pLSM-6 in *E. coli*.

Detailed Description and Preferred Embodiments

Disclosed herein is a method for the production of bioadhesive precursor protein analogs having incorporated therein hydroxyproline or other proline analogs. Vectors

comprising synthetic dsDNA sequences encoding the bioadhesive precursor protein analog are accordingly constructed and used to transform a microbial host, preferably *E. coli*. In an important feature , the transformed host is grown under conditions in which the protein is expressed, and subjected to osmotic shock in the presence of hydroxyproline (or other proline analog). This results in assimilation of hydroxyproline into the host, and incorporation of hydroxyproline into the bioadhesive precursor protein. There is thus no need to import any cellular machinery for post-translational hydroxylation of proline into the host or any need to perform other post-translational chemical or enzymatic proline hydroxylation. The resulting precursor protein is then isolated, purified if necessary, and the tyrosine residues hydroxylated by chemical or enzymatic means prior to cure.

While *E. coli* is presently preferred as a host, both prokaryotic and eukaryotic cells which can assimilate and incorporate proline analogs under hypertonic conditions may be used, including but not limited to prokaryotic bacteria, eukaryotic cells such as *Saccharomyces cervisiae*, *Saccharomyces carlsbergensis*, and *Schizosaccharomyces pombe*, and additional eukaryotes such as insect cells including lepidopteran cell lines such as *Spodoptera frugiperda*, *Trichoplasia ni*, *Heliothis virescens*, *Bombyx mori* infected with a baculovirus; CHO cells, COS cells, and NIH 3T3 cells.

The bioadhesive precursor protein analog produced by this method comprises from about 1% to about 50%, from about 5% to about 30%, or from about 10% to about 20% by total number of residues hydroxyproline, other proline analog, or mixtures thereof. Preferably, the bioadhesive precursor protein analog comprises a repeating decapeptide unit of the sequence

Ala-Lys-Pro/Hyp-Ser-Tyr-Pro/Hyp-Pro/Hyp-Thr-Tyr-Lys

wherein, within the bioadhesive precursor protein analog, at least one of the above residues designated Pro/Hyp is hydroxyproline, other protein analog, or mixtures thereof. The bioadhesive precursor protein analog preferably comprises from about three to about 100, from about three to about 50, or from about three to about twenty-five of the repeating decapeptide units of the above sequence. Another embodiment of the bioadhesive precursor protein comprises repeat units of the above decapeptide interspersed with at least one hexapeptide linker. One particularly useful hexapeptide linker has the sequence

Lys-Tyr-Pro/Hyp-Ser-Ala-Lys

From one to about thirty, one to about fifteen, or one to about five repeat DNA units encoding for the decapeptide may be interspersed between the hexapeptide linker. Other DNA sequences as disclosed by the prior art and encoding at least one proline residue may also be used for the production of a bioadhesive precursor protein having incorporated therein at least one hydroxyproline or other proline analog. In any one of the above embodiments, from about 1% to about 100%, from about 5% to about 75%, from about 15% to about 50%, or from about 5% to about 15% of the residues designated Pro/Hyp is hydroxyproline, other proline analog, or a combination thereof.

Hydroxyproline analogs of proline, such as *trans*-4-hydroxy-L-proline, *cis*-4-hydroxy-L-proline, *trans*-3-hydroxy-L-proline, *cis*-3-hydroxy-L-proline, and 3,4-dihydroxy-L-proline (and their corresponding D forms) are of particular interest in construction of a commercially successful bioadhesive precursor protein. However, other proline analogs which are imported into the host cell by the same mechanism may also be used, in order to determine

their effect on adhesion, to conduct structure/activity experiments, or the like. Such proline analogs include, but are not limited to, *cis*- and *trans*-4-fluoro-L-proline, *cis*- and *trans*-3-fluoro-L-proline, their D-analogs, 3,4-epoxyproline, and the like. Mixtures of two or more of any of the foregoing may also be used. For convenience, this specification uses the term "hydroxyproline" to refer inclusively to *trans*-4-hydroxy-L-proline, *cis*-4-hydroxy-L-proline, *trans*-3-hydroxy-L-proline, *cis*-3-hydroxy-L-proline, 3,4-dihydroxy-L-proline, their D-forms, and other proline analogs substituted with one or more hydroxy groups which are assimilated under hypertonic growth conditions and incorporated into proteins during translation. "Proline analogs" as used herein refers inclusively to the above and other analogs, as long as those analogs are assimilated into the host cells under hypertonic growth conditions, and incorporated into growing proteins at the translational stage by proline tRNA-synthetase.

Referring now to FIG. 2, the bioadhesive precursor protein analog is first produced by inserting a synthetic dsDNA sequence encoding the protein into a replicable expression vector. Within the vector, the sequence is operably linked to a regulatory sequence that is capable of directing expression of the encoded protein in *E. coli* (or other) host cells. The dsDNA is designed so as to encode for a precursor protein which is an analog of the naturally occurring *M. edulis* adhesive protein. As used herein "precursor protein analog" means a protein which may differ from the *M. edulis* protein in its exact amino acid sequence, but which comprises decapeptide repeating units which are common to the naturally occurring *M. edulis* protein except wherein the Dopa residues are substituted by tyrosine, and wherein at least one of the residues designated Pro/Hyp is hydroxyproline or other proline analog. The

dsDNA sequence is further designed so as to incorporate preferred *E. coli* codons. A preferred sequence is shown in FIG. 1.

The dsDNA sequence encoding the bioadhesive precursor protein decapeptide is prepared by any of a number of known methods of DNA synthesis. A suitable method for the synthesis of a dsDNA sequence is the methyl phosphite solid-phase method (Tetrahedron Letters, Vol. 21, pp. 719-722 (1980), and was used to synthesize the above-described dsDNA sequence using an automated solid-phase DNA synthesizer manufactured by Beckman in accordance with the manufacturer's directions. The DNA is then purified by preparative polyacrylamide gel electrophoresis. The purified DNA is enzymatically phosphorylated at the 5′-end prior to subsequent ligation.

These DNA sequences are then ligated to form multimer sequences having from three to about 100 repeats, from about three to about 50 decapeptide sequence repeats, from three to about 25 repeats, or as shown in FIG. 2, 10 repeats. These multimers are then purified and linker sequences providing unique restriction sites are added to the 5′ and 3′ ends to produce the sequence shown in FIG. 2. These sequences are then annealed, cleaved with the appropriate restriction site enzymes, and ligated into the restriction sites of an *E. coli* expression vector, i.e., pET-20b to generate pUSC-MAP, a plasmid that comprises ten of the decapeptide-coding multimers.

The same general procedure was used to construct pLSM-6 (FIG. 3) having 25 repeats of the DNA sequence encoding for a decapeptide having the sequence Ala-Lys-Pro/Hyp-Ser-Tyr-Pro/Hyp-Pro/Hyp-Thr-Tyr-Lys. pLSM-6 contains the bioadhesive precursor protein decapeptide as a C-terminal fusion to the periplasmic localization sequence

encoded on pET-20b. The 25-unit DNA insert was then subcloned into the expression vector pTrc99+ to produce a non-fusion clone pLSM-10 (FIG. 4).

A flow chart illustrating a general procedure for the construction of a plasmid encoding for a bioadhesive precursor protein having at least one interspersed polypeptide, in this instance a hexapeptide having the sequence Lys-Tyr-Pro/Hyp-Ser-Ala-Lys-CO_2H is shown in FIG. 5. Synthesis of DNA encoding for the decapeptide sequence -Ala-Lys-Pro/Hyp-Ser-Tyr-Pro/Hyp-Pro/Hyp-Thr-Tyr-Lys-CO_2H is followed by ligation into multimers having from one to about thirty, one to about fifteen, or one to about five DNA repeats, and purification of the ligated DNA. Linkers encoding for restriction sites (e.g., EcoR I, Nco I, DdeI, Hind III, BamH I, and Xba I) are then ligated to the multimers by standard techniques, and the resulting DNA is purified.

Importantly, the restriction site for Dde I is ligated immediately adjacent each end of the multimer encoding for the bioadhesive precursor protein. Treating this DNA with the restriction enzyme Dde I results in DNA which can then be self-ligated using the cut ends. Self-ligation results in DNA sequences having from one to about thirty, one to about fifteen, or one to about five DNA sequences encoding for the decapeptide multimers interspersed with a hexapeptide unit having the sequence Lys-Tyr-Pro/Hyp-Ser-Ala-Ly. Linkers suitable for cloning into an expression vector are then ligated, and the DNA is used to transform an appropriate vector. This method allows the assembly of greater numbers of multimers of the DNA units encoding for bioadhesive protein precursor. Other DNA/amino acid sequence/restriction enzyme combinations (as embodied by the above-described hexapeptide sequence/Dde I combination) may be used as long as the resultant amino acid sequence does

not inhibit protein expression, isolation, adhesiveness, or other desired bioadhesive precursor protein properties. Amino acid sequences may be up to fifty, and even greater number of residues in length. However, shorter sequences are preferred for reasons of economy and specificity.

The expression vector containing the inserted multimers coding for the bioadhesive precursor protein analog and optional hexapeptide is then used to transform *E. coli* cells by known techniques of transformation. The transformed *E. coli* cells are cultured under conditions suitable for growth and expression of the bioadhesive precursor protein analog gene. Incorporation of hydroxyproline into proteins during protein synthesis has been disclosed in commonly-assigned U.S. Pat. Application Ser. No. 08/655,086 filed June 3, 1996 to Gruskin et al., which is incorporated by reference herein in its entirety. As disclosed therein, the normal process of polypeptide or protein synthesis in living cells requires the transcription of DNA into RNA, and translation of RNA into protein. DNA is first transcribed into messenger RNA (mRNA), which carries the genetic information to ribosomes made of ribosomal RNA (rRNA). Transfer RNA (tRNA) meanwhile links to free amino acids in the cell pool to form amino acid/tRNA complexes. The amino acid/tRNA complexes then align with the codons of mRNA, with actual recognition and binding being mediated by tRNA. Each of up to about twenty amino acids is distinguished from other amino acids and charged to tRNA by enzymes known as aminoacyl-tRNA synthetases. Amino acid/tRNA complexes are generally quite specific, such that normally only a molecule with an exact stereochemical configuration is acted upon by a particular aminoacyl-tRNA synthetase.

Certain rare exceptions exist however, including the amino acid activating enzymes responsible for the synthesis of the prolyl-tRNA complex. This feature of *E. coli* (and certain other prokaryotes and eukaryotes) can be exploited under certain conditions to allow the synthesis of proteins containing proline analogs such as hydroxyproline. Thus, a cell capable of producing the desired protein, when contacted with a preferably hypertonic growth medium comprising hydroxyproline, assimilates the hydroxyproline into the cell and incorporates it into an amino acid/tRNA complex, thereby producing proteins having hydroxyproline in place of proline. This assimilation and incorporation is extremely useful when the structure and function of a polypeptide or protein depends on post-translational hydroxylation of proline not present in the native production system of a recombinant host, such as *E. coli, Saccharomyces cervisiae*, and eukaryotic cell lines including *Spodoptera frugiperda, Tricholplasia, ni, Heliothis virescens,*, and others.

In the present instance, a culture of the transformed *E. coli* strain is grown in a rich (i.e., LB) media until an OD_{600} of about 0.75 - 1.00 is reached. The cells are harvested, washed with a minimal (i.e., M9) media supplemented with the 20 natural amino acids minus proline, and the media replaced by an equal amount of the identical minimal medium except for the addition of 400 mM NaCl. After incubation, the cells are induced (in the present instance by addition of isopropyl-β-D-thiogalactopyranoside) and the proline analog is simultaneously added. The cells are then incubated for an additional length of time.

An advantageous feature of the present invention is that the proline analog content may be adjusted and optimized by the nature and concentration of proline analogs added at induction of the host cells. Thus, the total hydroxyproline content of the bioadhesive

precursor protein may be adjusted by the addition of suitable mixtures of proline and hydroxyproline. Use of such mixtures is readily empirically determined by one of ordinary skill in the art, and allows the production of bioadhesive precursor proteins having from 1% to 50%, 5% to 30%, or 10% to 20% hydroxyproline or other proline analogs, depending on the bioadhesive precursor protein amino acid sequence. Treatment with a mixture of proline analogs may result in a protein having incorporated therein a mixture of the protein analogs. Mixtures of proline analogs suitable for producing such proteins is empirically determined, and will depend on the relative rates of analog uptake, aminoacyl t-RNA charging, and incorporation into the protein during translation. Novel proteins are available using this method, in that the naturally-occurring bioadhesive proteins are known to comprise a mixture of proline, *trans*-4-hydroxy-L-proline, 3-hydroxy-L-proline, and sometimes 3,4-dihydroxy-L-proline. Use of a treatment mixture comprising exclusively *trans*-4-hydroxy-L-proline (with or without proline), for example, yields a bioadhesive precursor protein wherein the Pro/Hyp residues consist of proline and from 1% up to 100% *trans*-4-hydroxy-L-proline but no other proline analogs. Use of a treatment mixture comprising exclusively 3-hydroxy-L-proline (with or without proline) yields a bioadhesive precursor protein wherein the Pro/Hyp residues consist of proline and from 1% up to 100% 3-hydroxy-L-proline and no other proline analogs. These proteins are novel, and may provide useful information regarding the structure and function of bioadhesive precursor proteins.

After the protein has been expressed, it is recovered from the transformant cells by known methods such as mechanical or chemical lysis of the cells. The protein is preferably purified, for example by dialysis and chromatographic procedures well-known in the art to

homogeneity or near-homogeneity. In the case of a fusion protein, the recovered protein is preferably subjected to cyanogen bromide cleavage to remove extraneous peptide sequences.

The bioadhesive precursor protein analog is converted to adhesive by methods known in the art, for example by chemical or enzymatic hydroxylation of tyrosine residues and subsequent oxidation of the resultant DOPA residues. Enzymatic hydroxylation of tyrosine may be effected in the presence of mushroom tyrosinase as described by Ito et al. in *Biochemistry*, vol. 22, pp. 407-411 (1984) and Marumo and Waite in *Biochem. and Biophys. Acta*, Vol. 892, pp. 98-103 (1986).

The following examples are included for purposes of illustration and are not to be construed as limitations herein. Amino acid sequences as used herein follow the convention wherein the first (left-most) amino acid is the N-terminal amino acid and the last (right-most) amino acid is the C-terminal amino acid.

EXAMPLE 1

Synthesis and Cloning of a Gene Encoding a Mussel Adhesive Protein Analog

A DNA sequence was synthesized using codon preferences favorable for expression in *E. coli*. as follows:

5′-GCT AAA CCG TCC TAC CCA CCG ACC TAC AAG
CGA TTT GGC AGG ATG GGT GGC TGG ATG TTC

which sequence corresponds to the mussel adhesive protein analog decamer amino acid sequence

Ala-Lys-Pro/Hyp-Ser-Tyr-Pro/Hyp-Pro/Hyp-Thr-Tyr-Lys

(hereinafter the "MAP analog") using a Beckman Oligo 1000M DNA synthesizer, and purified by preparative gel electrophoresis. The two DNA strands were synthesized to contain a three base (one codon) overhang following annealing as shown in FIG. 2. Each DNA oligomer (800 pmol) was 5'-phosphorylated in a 25 μL reaction mixture containing T4 polynucleotide kinase and 1 mM ATP (37 °C, 2.5 hrs). The top and bottom strands were annealed by mixing, heating to 80 °C for 15 minutes, and allowing to slow-cool to 30°C, followed by self-ligation via the addition of 1 mM ATP, T4 DNA ligase buffer, and T4 DNA ligase (60 μL total) and incubation at 16 °C for 16 hours. The ligated multimers were then capped by the addition of synthetic linker oligonucleotide duplexes (200 pmol) containing restriction sites for cloning. The ligation reaction was allowed to continue for an additional 2 hours and the ligase inactivated by heating at 65 °C for 20 min and slow cooling to 30 °C.

The mixture of capped oligonucleotides encoding the MAP analog was cleaved with *Nco I* and *Hind III* (37 °C, 3 hours), gel purified on a 2% agarose gel (1X TPE buffer, 1.5 hours, 60 V), and ligated into the *Nco I* and *Hind III* restriction sites of expression vector pET-20b to yield PLSM-6 (FIG. 3). Clones containing various lengths of polymeric sequence were selected from a population of transformant clones containing the MAP analog DNA insert. A clone containing 25 repeats of the MAP analog DNA sequence was verified by standard sequencing methods. The MAP analog DNA insert was subsequently subcloned into the expression vector pTrc99+ at the *Nco I* and *Hind III* sites to produce a non-fusion clone pLSM-10 (FIG. 4).

EXAMPLE 2

Expression of MAP Analog Polypeptide Sequence and Incorporation of *trans*-4-Hydroxy-L-Proline

Plasmid pLSM-10 obtained in Example 1 (FIG. 4) was used to transform proline auxotrophic *E. coli* strain JM109F-(pro$^-$). A 20 mL culture of the transformed *E. coli* strain was grown in LB medium until an OD_{600} of 0.75 - 1.00 was reached. The cells were harvested by centrifugation, washed with M9 minimal medium (M9 salts, 2% glucose, 0.01 mg/mL thiamine, supplemented with the 20 natural amino acids minus proline at 50 μg/mL), and the medium replaced by an equal amount of the identical M9 minimal medium with the additon of 400 mM NaCl. After incubation for 30 minutes at 37 °C, the cells were induced by addition of isopropyl-β-D-thiogalactopyranoside (IPTG) to a concentration of 1.5 mM and *trans*-4-hydroxy-L-proline was added to provide a final concentration of 15 mM. The cells were incubated for an additional 4 hours at 37 °C with vigorous shaking (250 rpm). The cells were harvested by centrifugation and lysed by sonication (3 times for 20 seconds each). An aliquot of both the soluble and insoluble cell fractions were electrophoresed on an 12% SDS-PAGE gel using control lanes. As shown in FIG. 6, the analysis indicates the presence of an inducible protein band migrating between 30-40 kDa.

EXAMPLE 3

Purification of the Expressed MAP Analog Polypeptide Comprising *trans*-4-Hydroxy-L-Proline

The cell pellet from Example 2 is resuspended in one-fiftieth of the volume of the original culture in lysis buffer (50 mM sodium phosphate, pH 6.5, 100 mM NaCl, 10 mM

EDTA, 5 mM DTT) and lysed by standard methods (French press or sonication). Protein in the insoluble cell fraction is solubilized by addition of 0.8 M acetic acid, 2 M urea, and dialyzed first against 0.5 M acetic acid (2 changes, 3 hours each, 4 °C) and lyophilized to dryness. The dry protein material is redissolved in a suitable volume of 0.1 M HCl, and a 5 to 50-fold molar excess of cyanogen bromide is added. The reaction is allowed to stir in the dark at room temperature for about 24 hours. The solvent is then removed by evaporation, and the bioadhesive precursor protein is further purified by standard chromatographic methods such as reverse-phase (e.g., C-4) high pressure liquid chromatography (HPLC) or cation exchange chromatography (e.g., CM cellulose), dialyzed against 0.5 M acetic acid, and lyophilized to yield purified protein comprising the MAP analog wherein approximately 95-100% of the Hyp/Pro residues are *trans*-4-hydroxy-L-proline.

EXAMPLE 4

Cure of a Protein Comprising the MAP Analog Polypeptide Comprising *trans*-4-Hydroxy-L-Proline

The protein of Example 3 is concentrated to 0.8 mg/mL in 0.1 M phosphate buffer adjusted to pH 7.0 with 25 mM ascorbic acid. Tyrosinase (obtained from Sigma) and used as received) (1 unit/mL) is added and the solution is incubated at 37 °C for 1 hour (thereby oxidizing at least some tyrosine residues to dopamine residues), followed by dialysis against 10% acetic acid, lyophilization, and storage at -20 °C. The protein may also be stored in 5% acetic acid or in borate buffer adjusted to pH 8.5. Complete cure (i.e, cross-linking of the dopamine to produce a cross-linked gel or to attach a substrate) of the protein obtained in

Example 3 or the stored dopamine-containing protein after tyrosinase incubation is effected by incubation with tyrosinase at neutral pH in the absence of ascorbate.

EXAMPLE 5

Expression of the MAP Analog Polypeptide Incorporating Proline and *trans*-4-Hydroxy-L-Proline

Plasmid LSM-6 containing DNA encoding 25 repeats of the MAP analog sequence (FIG. 3) was used to transform proline auxotrophic *E. coli* strain JM109(DE3)pLysS (pro$^-$). A 20 mL culture of the transformed *E. coli* strain was grown in LB medium until an OD_{600} of about 1.0 was reached. The cells were harvested by centrifugation, washed with M9 minimal media (M9 salts, 2% glucose, 0.01 mg/mL thiamine, 200 µg/mL chloramphenicol, supplemented with the 20 natural amino acids minus proline at 50 µg/mL), and the media replaced by an equal amount of the same M9 minimal medium except for the presence of 400 mM NaCl. After 30 minutes equilibration at 37 °C, the cells were induced by addition of isopropyl-β-D-thiogalactopyranoside (IPTG) to a concentration of 1.5 mM, *trans*-4-hydroxy-L-proline to a concentration of 15 mM, and L-proline to a concentration of 10 µM. The cells were incubated for an additional 4 hours at 37°C with vigorous shaking (250 rpm). The cells were harvested by centrifugation, and lysed by sonication in lysis buffer (50 mM sodium phosphate buffer, pH 6.4, 10 mM EDTA, 100 mM NaCl, 5 mM 2-mercaptoethanol, and 250 µg/mL lysozyme). Soluble and insoluble cell fractions from one experiment are shown in FIG. 6, and the insoluble cell fraction from another experiment in FIG. 7 at lane 1. All MAP proteins appear in the insoluble fraction by SDS-PAGE at 30-40 kDa.

EXAMPLE 6

Purification of the Expressed MAP Analog Polypeptide Comprising Proline and *trans*-4-Hydroxy-L-Proline

The cell pellet from Example 5 is resuspended in one-fiftieth of the volume of the original culture in lysis buffer (50 mM sodium phosphate, pH 6.5, 100 mM NaCl, 10 mM EDTA, 5 mM DTT) and lysed by standard methods (French press or sonication). Protein in the insoluble cell fraction is solubilized by addition of 0.8 M acetic acid, 2 M urea, and dialyzed first against 0.5 M acetic acid (2 changes, 3 hours each, 4 °C) and lyophilized to dryness. The dry protein material is redissolved in a suitable volume of 0.1 M HCl, and a 5 to 50-fold molar excess of cyanogen bromide is added. The reaction is allowed to stir in the dark at room temperature for about 24 hours. The solvent is then removed by evaporation, and the bioadhesive precursor protein I further purified by standard chromatographic methods such as reverse-phase (e.g., C-4) high pressure liquid chromatography (HPLC) or cation exchange chromatography (e.g., CM cellulose), dialyzed against 0.5 M acetic acid, and lyophilized to yield purified protein comprising the MAP analog having incorporated therein *trans*-4-hydroxy-L-proline and proline.

EXAMPLE 7.

Curing of the MAP Analog Polypeptide With Incorporated Proline and *trans*-4-Hydroxy-L-Proline

The MAP analog protein with incorporated proline analogue expressed as in Example 5 and purified as in Example 6 is concentrated to 0.8 mg/mL in 0.1 M phosphate buffer and 25 mM ascorbic acid, adjusted to pH 7.0. Tyrosinase (1 unit/mL) is added and the reaction

incubated at 37 °C for 1 hours.(thereby oxidizing at least some tyrosine residues to dopamine residues), followed by dialysis against 10% acetic acid, lyophilization, and storage at -20 °C. The protein may also be stored in 5% acetic acid or in borate buffer adjusted to pH 8.5. Complete cure (i.e, cross-linking of the dopamine to produce a cross-linked gel or to attach a substrate) of the protein obtained in Example 6 or the stored dopamine-containing protein after tyrosinase incubation is effected by incubation with tyrosinase at neutral pH in the absence of ascorbate.

EXAMPLE 8

Expression of the MAP Analog Polypeptide Incorporating Proline and *cis*-4-Hydroxy-L-Proline

Plasmid LSM-6 containing DNA encoding 25 repeats of the MAP analog sequence (FIG. 3) was used to transform proline auxotrophic *E. coli* strain JM109(DE3)pLysS (pro^-). A 20 mL culture of the transformed *E. coli* strain was grown in LB medium until an OD_{600} of about 1.0 was reached. The cells were harvested by centrifugation, washed with M9 minimal media (M9 salts, 2% glucose, 0.01 mg/mL thiamine, 200 µg/mL chloramphenicol, supplemented with the 20 natural amino acids minus proline at 50 µg/mL), and the media replaced by an equal amount of the same M9 minimal medium except for the presence of 400 mM NaCl. After 30 minutes equilibration at 37 °C, the cells were induced by addition of isopropyl-β-D-thiogalactopyranoside (IPTG) to a concentration of 1.5 mM, *cis*-4-hydroxy-L-proline to a concentration of 15 mM, and proline to a concentration of 10 µM. The cells were incubated for an additional 4 hours at 37°C with vigorous shaking (250 rpm). The cells

were harvested by centrifugation, and lysed by sonication in lysis buffer (50 mM sodium phosphate buffer, pH 6.4, 10 mM EDTA, 100 mM NaCl, 5 mM 2-mercaptoethanol, and 250 μg/mL lysozyme). The insoluble cell fraction is shown in FIG. 7, lane 2. All MAP proteins appear in the insoluble fraction by SDS-PAGE at 30-40 kDa.

EXAMPLE 9

Purification of the Expressed MAP Analog Polypeptide With Incorporated Proline and *cis*-4-Hydroxy-L-Proline

The cell pellet from Example 8 is resuspended in one-fiftieth of the volume of the original culture in lysis buffer (50 mM sodium phosphate, pH 6.5, 100 mM NaCl, 10 mM EDTA, 5 mM DTT) and lysed by standard methods (French press or sonication). Protein in the insoluble cell fraction is solubilized by addition of 0.8 M acetic acid, 2 M urea, and dialyzed first against 0.5 M acetic acid (2 changes, 3 hours each, 4 °C) and lyophilized to dryness. The dry protein material is redissolved in a suitable volume of 0.1 M HCl, and a 5 to 50-fold molar excess of cyanogen bromide is added. The reaction is allowed to stir in the dark at room temperature for about 24 hours. The solvent is then removed by evaporation, and the bioadhesive precursor protein I further purified by standard chromatographic methods such as reverse-phase (e.g., C-4) high pressure liquid chromatography (HPLC) or cation exchange chromatography (e.g., CM cellulose), dialyzed against 0.5 M acetic acid, and lyophilized to yield purified protein comprising the MAP analog having incorporated therein proline and *cis*-4-hydroxy-L-proline.

EXAMPLE 10

Curing of the MAP Analog Polypeptide With Incorporated Proline and *cis*-4-Hydroxy-L-Proline

The MAP analog protein with incorporated proline analogue expressed as in Example 8 and purified as in Example 9 is concentrated to 0.8 mg/mL in 0.1 M phosphate buffer and 25 mM ascorbic acid, adjusted to pH 7.0. Tyrosinase (1 unit/mL) is added and the reaction incubated at 37 °C for 1 hours.(thereby oxidizing at least some tyrosine residues to dopamine residues), followed by dialysis against 10% acetic acid, lyophilization, and storage at -20 °C. The protein may also be stored in 5% acetic acid or in borate buffer adjusted to pH 8.5. Complete cure (i.e, cross-linking of the dopamine to produce a cross-linked gel or to attach a substrate) of the protein obtained in Example 6 or the stored dopamine-containing protein after tyrosinase incubation is effected by incubation with tyrosinase at neutral pH in the absence of ascorbate.

While preferred embodiments have been shown and described, various modifications and substitutions may be made thereto without departing from the spirit and scope of the invention. Accordingly, it is to be understood that the invention has been described by way of illustration and not limitation.

What is claimed is:

1. A method for producing a bioadhesive precursor protein analog having a proline analog incorporated therein, comprising

constructing a DNA sequence comprising DNA encoding a bioadhesive precursor protein analog, said DNA sequence having at least one codon encoding proline;

inserting the DNA sequence into an expression vector capable of effecting expression of said protein in a host cell;

transforming host cells with the expression vector; and

expressing the bioadhesive precursor protein analog by contacting the transformed host cells with hypertonic growth media comprising the proline analog, thereby assimilating the proline analog into the host cells, and incorporating the proline analog into the bioadhesive precursor protein analog.

2. The method of claim 1, wherein the bioadhesive precursor protein analog comprises from about 1% to about 50% of one or more proline analog residues based on the total number of amino acid residues in the protein.

3. The method of claim 2, wherein the bioadhesive precursor protein analog comprises from about 3% to about 30% of one or more proline analog residues based on the total number of amino acid residues in the protein.

4. The method of claim 2, wherein the bioadhesive precursor protein analog comprises from about 5% to about 20% of one or more proline analog residues based on the total number of amino acid residues in the protein.

5. The method of claim 1, wherein the proline analog is selected from the group consisting of *hydroxyproline, trans*-4-hydroxy-L-proline, *cis*-4-hydroxy-L-proline, *trans*-3-hydroxy-L-proline, *cis*-3-hydroxy-L-proline, 3,4-dihydroxy-L-proline, *cis*-4-fluoro-L-proline, *trans*-4-fluoro-L-proline, *cis*-3-fluoro-L-proline, *trans*-3-fluoro-L-proline, 3,4-epoxyproline, and combinations thereof.

6. The method of claim 5, wherein the proline analog is hydroxyproline.

7. The method of claim 6, wherein the proline analog is *trans*-4-hydroxy-L-proline.

8. The method of claim 1, wherein the DNA sequence comprises at least three repeat units of the sequence 5'-GCTAAACCGTCCTACCCACCGACCTACAAG.

9. The method of claim 8, wherein the DNA sequence comprises from three to about 100 of the repeat units.

10. The method of claim 8, wherein the DNA sequence comprises from three to about fifty of the repeat units.

11. The method of claim 8, wherein the DNA sequence comprises from three to about twenty-five of the repeat units.

12. The method of claim 1, wherein the bioadhesive precursor protein analog comprises at least one decapeptide having the sequence

Ala-Lys-Pro/Hyp-Ser-Tyr-Pro/Hyp-Pro/Hyp-Thr-Tyr-Lys

wherein at least one of the residues designated Pro/Hyp is the one or more proline analogs.

13. The method of claim 12, wherein from about 1% to about 100% of the Pro/Hyp residues of the bioadhesive precursor protein analog is one or more proline analogs.

14. The method of claim 12, wherein up from about 5% to about 75% of the Pro/Hyp residues of the bioadhesive precursor protein analog is one or more proline analogs.

15. The method of claim 12, wherein from about 15% about 50% of the Pro/Hyp residues of the bioadhesive precursor protein analog is one or more proline analogs.

16. The method of claim 12, wherein from about 5% to about 15% of the Pro/Hyp residues of the bioadhesive precursor protein analog is one or more proline analogs.

17. The method of claim 12, wherein the proline analog is selected from the group consisting of hydroxyproline, *trans*-4-hydroxy-L-proline, *cis*-4-hydroxy-L-proline, *trans*-3-hydroxy-L-proline, *cis*-3-hydroxy-L-proline, 3,4-dihydroxy-L-proline, *cis*-4-fluoro-L-proline, *trans*-4-fluoro-L-proline, *cis*-3-fluoro-L-proline, *trans*-3-fluoro-L-proline, 3,4-epoxyproline, and combinations thereof.

18. The method of claim 17, wherein the proline analog is hydroxyproline.

19. The method of claim 18, wherein the proline analog is *trans*-4-hydroxy-L-proline.

20. The method of claim 8, wherein the DNA units are interspersed with at least one second DNA unit having a sequence encoding for a polypeptide from about four to about 12 amino acid residues, and further wherein the DNA sequence is cleavable by at least one restriction enzyme.

21. The method of claim 20, wherein the second DNA unit has the sequence 5'-AAATACCCCTCAGCTAAA.

22. The method of claim 12, wherein at least one of the decapeptides is adjacent a hexapeptide having the sequence Lys-Tyr-Pro/Hyp-Ser-Ala-Lys.

23. The method of claim 22, wherein at least one hexapeptide sequence is interspersed between multimers of the decapeptide sequence, wherein the multimers have from one to about thirty, one to about fifteen, or one to about five repeat decapeptide units.

24. A bioadhesive precursor protein analog produced by the method of claim 1, wherein up to about 75% of the Pro/Hyp residues of the bioadhesive precursor protein analog is one or more proline analogs

25. A bioadhesive precursor protein analog comprising the amino acid sequence

Ala-Lys-Pro/Hyp-Ser-Tyr-Pro/Hyp-Pro/Hyp-Thr-Tyr-Lys

wherein the Pro/Hyp residues consist of proline and from about 1% up to about 100% *trans*-4-hydroxy-L-proline.

26. A bioadhesive precursor protein analog comprising the amino acid sequence

Ala-Lys-Pro/Hyp-Ser-Tyr-Pro/Hyp-Pro/Hyp-Thr-Tyr-Lys

wherein the Pro/Hyp residues consist of proline and from about 1% up to about 100% *cis*-4-hydroxy-L-proline.

27. A bioadhesive precursor protein analog comprising the amino acid sequence

Ala-Lys-Pro/Hyp-Ser-Tyr-Pro/Hyp-Pro/Hyp-Thr-Tyr-Lys

wherein the Pro/Hyp residues consist of proline and from about 1% up to about 100% *trans*-3-hydroxy-L-proline.

28. A bioadhesive precursor protein analog comprising the amino acid sequence

Ala-Lys-Pro/Hyp-Ser-Tyr-Pro/Hyp-Pro/Hyp-Thr-Tyr-Lys

wherein the Pro/Hyp residues consist of proline and from about 1% up to about 100% *cis*-3-hydroxy-L-proline.

FIG_1

Ala Lys Pro Ser Tyr Pro Pro Thr Tyr Lys
GCT AAA CCG TCC TAC CCA CCG ACC TAC AAG
CGA TTT GGC AGG ATG GGT GGC TGG ATG TTC

FIG_3

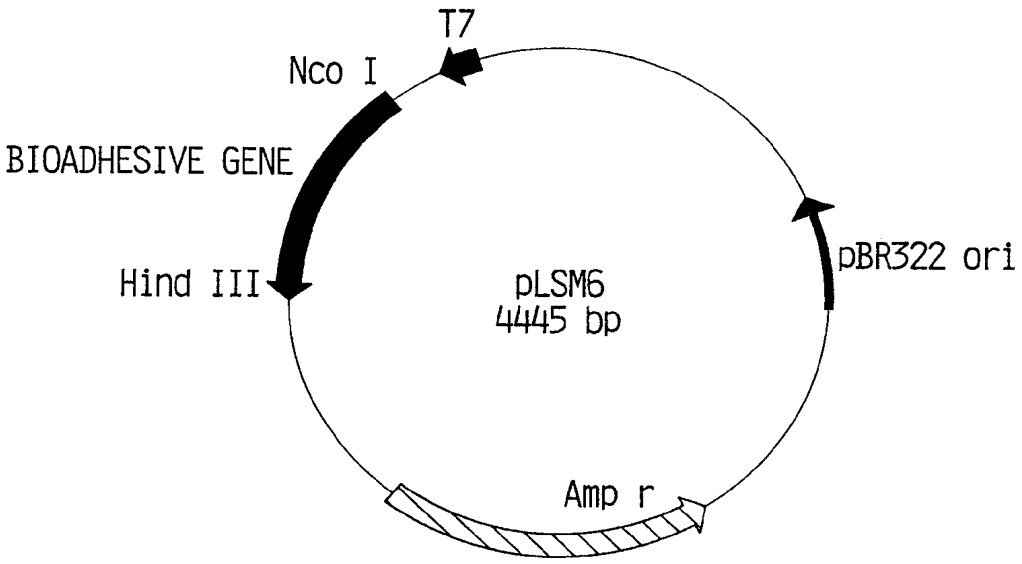


FIG_4

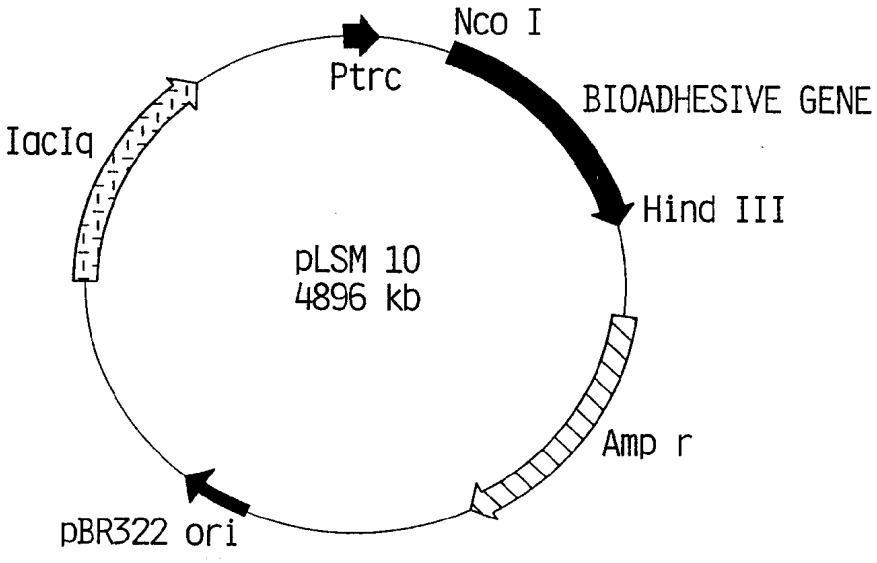


FIG_2

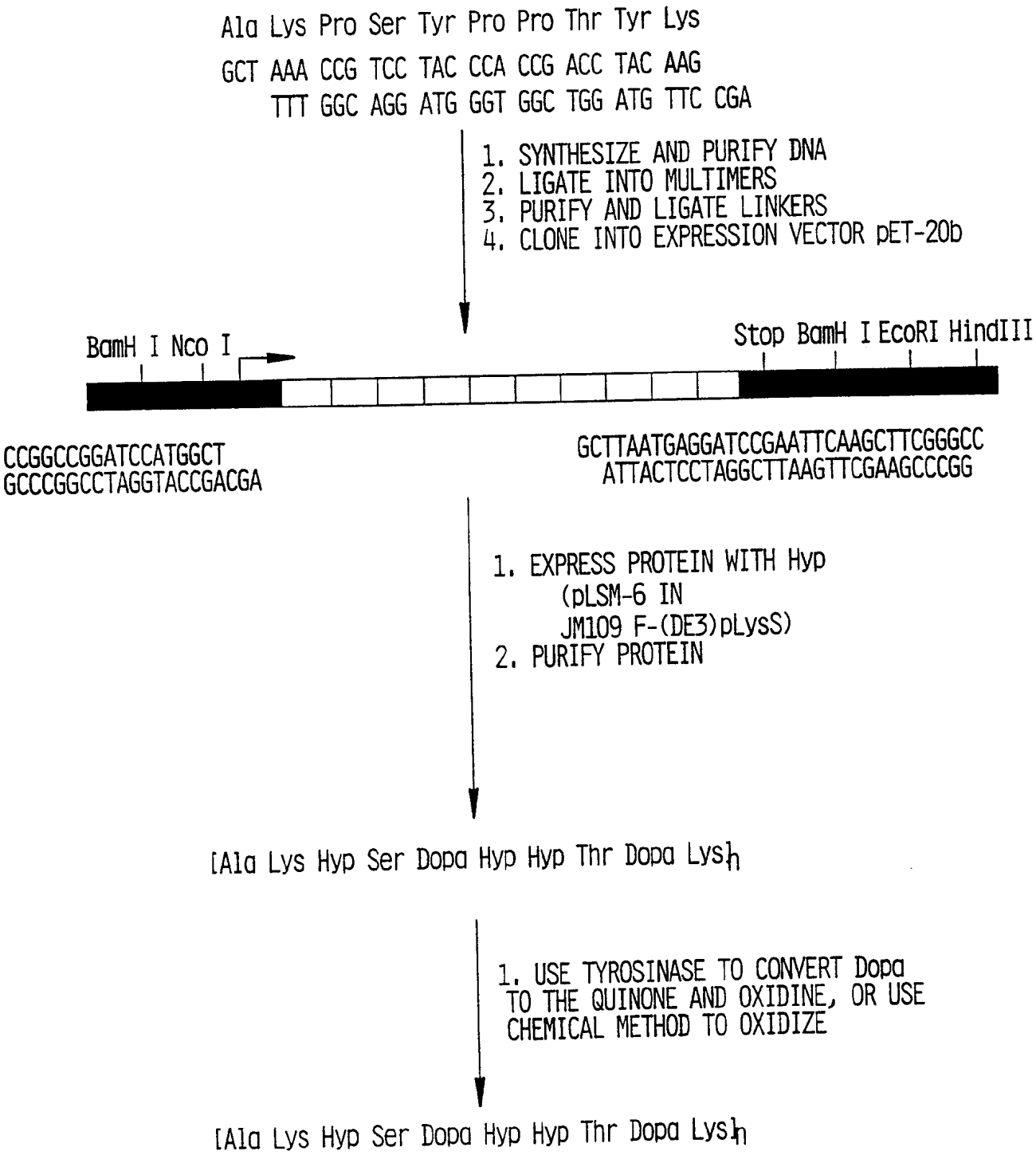

Ala Lys Pro Ser Tyr Pro Pro Thr Tyr Lys
GCT AAA CCG TCC TAC CCA CCG ACC TAC AAG
TTT GGC AGG ATG GGT GGC TGG ATG TTC CGA
1. SYNTHESIZE AND PURIFY DNA
2. LIGATE INTO MULTIMERS
3. PURIFY AND LIGATE LINKERS
4. CLONE INTO EXPRESSION VECTOR pET-20b
BamH I Nco I
Stop BamH I EcoRI HindIII
CCGGCCGGATCCATGGCT
GCCCGGCCTAGGTACCGACGA
GCTTAATGAGGATCCGAATTCAAGCTTCGGGCC
ATTACTCCTAGGCTTAAGTTCGAAGCCCGG
1. EXPRESS PROTEIN WITH Hyp
(pLSM-6 IN
JM109 F-(DE3)pLysS)
2. PURIFY PROTEIN
[Ala Lys Hyp Ser Dopa Hyp Hyp Thr Dopa Lys]n
1. USE TYROSINASE TO CONVERT Dopa
TO THE QUINONE AND OXIDINE, OR USE
CHEMICAL METHOD TO OXIDIZE
[Ala Lys Hyp Ser Dopa Hyp Hyp Thr Dopa Lys]n

FIG_5

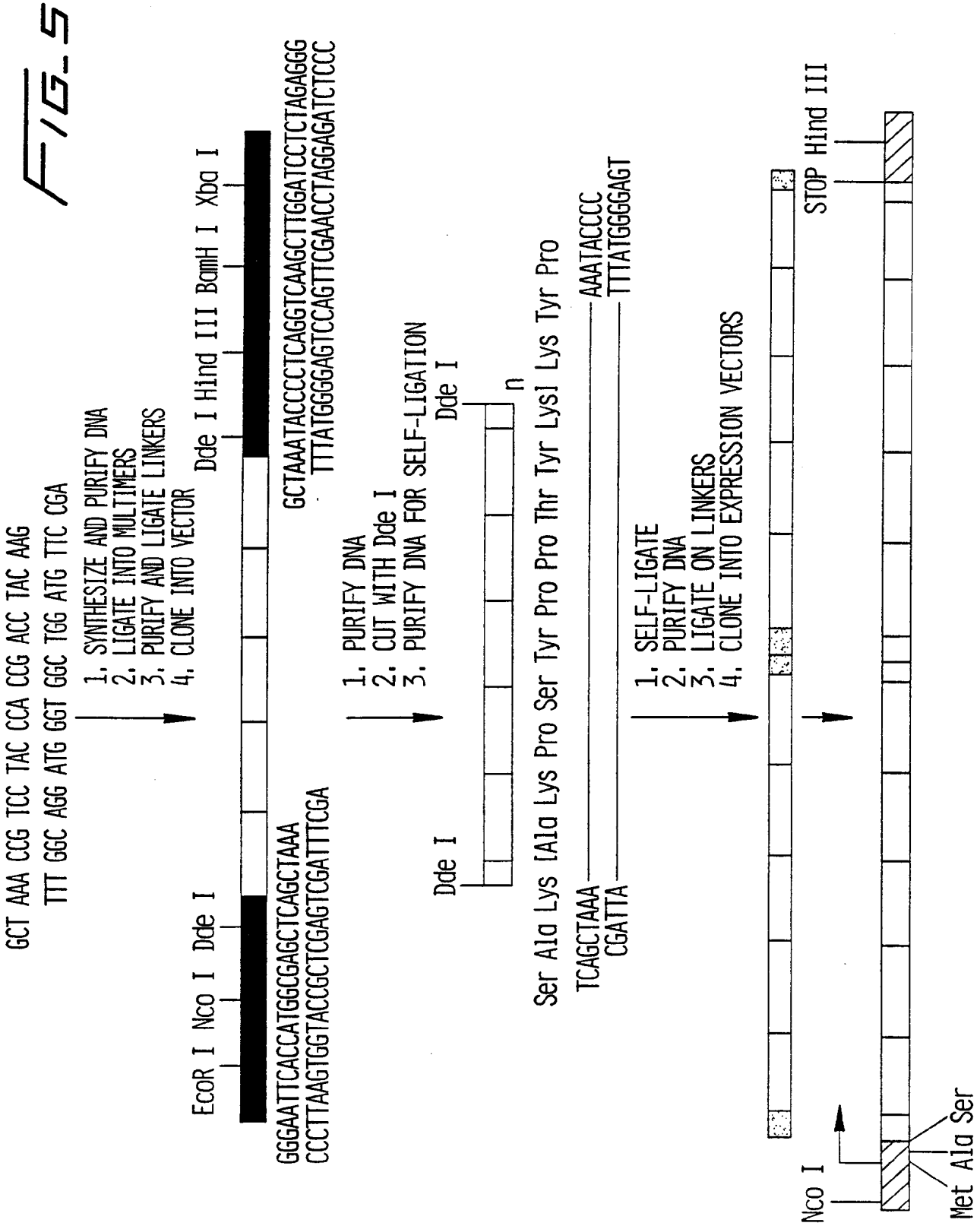

GCT AAA CCG TCC TAC CCA CCG ACC TAC AAG
TTT GGC AGG ATG GGT GGC TGG ATG TTC CGA
1. SYNTHESIZE AND PURIFY DNA
2. LIGATE INTO MULTIMERS
3. PURIFY AND LIGATE LINKERS
4. CLONE INTO VECTOR
EcoR I Nco I Dde I
Dde I Hind III BamH I Xba I
GGGAATTCACCATGGCGAGCTCAGCTAAA
CCCTTAAGTGGTACCGCTCGAGTCGATTTCGA
GCTAAATACCCCTCAGGTCAAGCTTGGATCCTCTAGAGGG
TTTATGGGGAGTCCAGTTCGAACCTAGGAGATCTCCC
1. PURIFY DNA
2. CUT WITH Dde I
3. PURIFY DNA FOR SELF-LIGATION
Dde I
n
Dde I
Ser Ala Lys [Ala Lys Pro Ser Tyr Pro Pro Thr Tyr Lys] Lys Tyr Pro
TCAGCTAAA
CGATTA
AAATACCCC
TTTATGGGGAGT
1. SELF-LIGATE
2. PURIFY DNA
3. LIGATE ON LINKERS
4. CLONE INTO EXPRESSION VECTORS
STOP Hind III
Nco I
Met Ala Ser

FIG_6

	CELL LYSATE, SOLUBLE					CELL LYSATE, INSOLUBLE				
Hyp(15mM)	X	X	X			X	X	X		
Pro(10μM)	X	X		X		X	X		X	
IPTG(1.5mM)	X		X	X	X	X		X	X	X
NaCl(400mM)	X	X	X	X	X	X	X	X	X	X

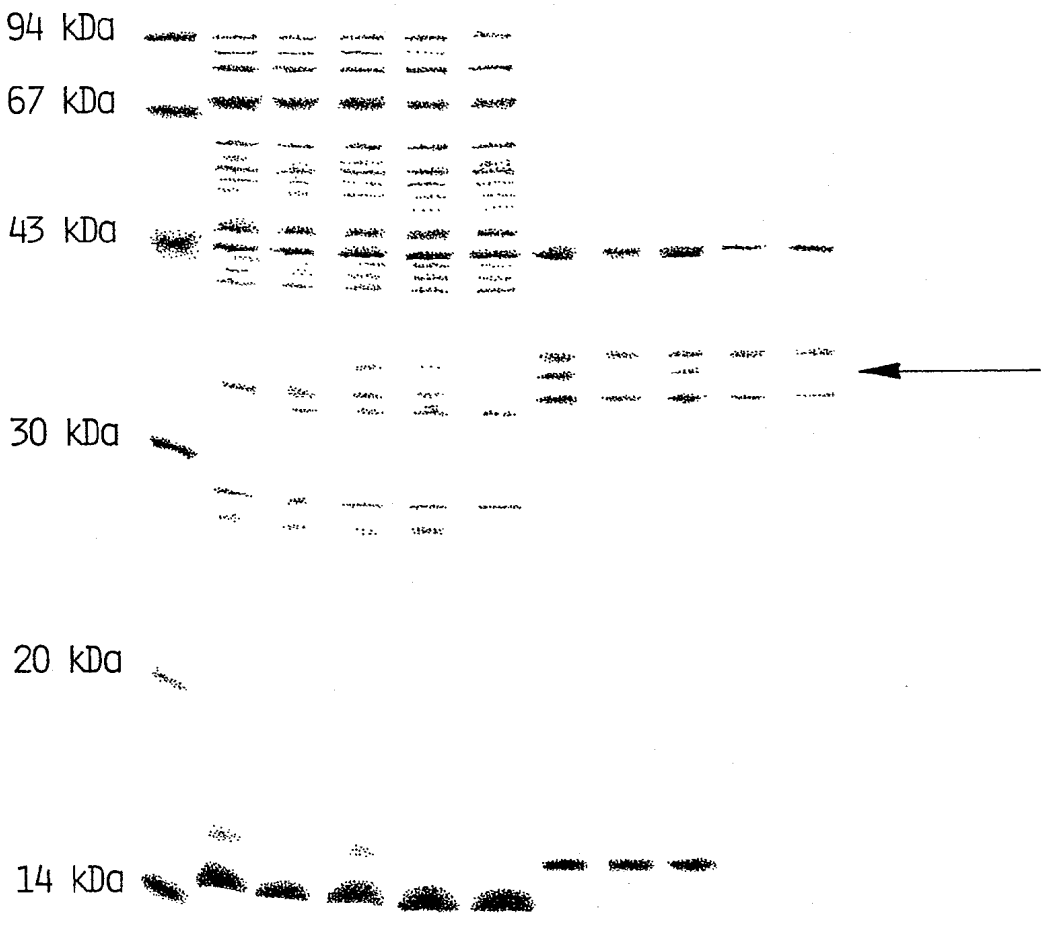
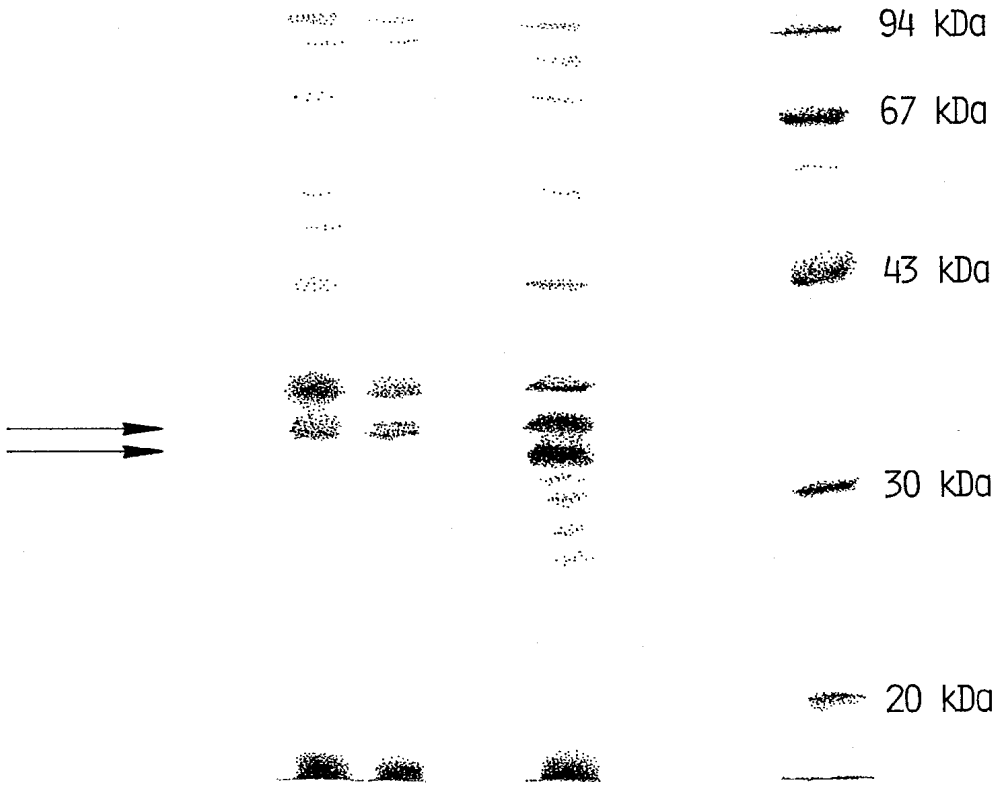


FIG. 7

Pro(1mM)				X		
cis-4-Hyp(15mM)		X				
Hyp(15mM)	X					
Pro(10μM)	X	X				
IPTG(1.5mM)	X	X		X		



<110> Paolella, David N.
Gruskin, Elliott A.
Buechter, Douglas D.

<120> Bioadhesive Protein Analogs Comprising Hydroxyproline and Methods of Production Thereof

<130> 98-1510

<140> US 09/149,973
<141> 1998-09-09

<160> 15

<210> 1
<211> 10
<212> PRT
<213> Mytilus edulis

<220>
<221> VARIANT
<222> (3)..(9)
<223> Positions 3, 6 and 7 are proline or hydroxyproline residues. Positions 4 and 8 are serine or threonine residues. Positions 5 and 9 are tyrosine or dopamine residues.

<400> 1
Ala Lys Xaa Xaa Xaa Xaa Xaa Xaa Xaa Lys
1 5 10

<210> 2
<211> 8
<212> PRT
<213> G. demissa

<220>
<221> VARIANT
<222> (1)
<223> Position 1 is a Thr or Ala residue.

<220>
<221> VARIANT
<222> (3)..(5)

<223> Position 3 is a Dopamine or Tyr residue. Position 4 is a Ser, Val, Asp, or Leu residue. Position 5 is an Ala, Pro, Hyrdoxyproline, or Leu residue.

<220>
<221> VARIANT
<222> (7)

<223> Position 7 is a Dopamine or Tyr residue.

<400> 2
Xaa Gly Xaa Xaa Xaa Gly Xaa Lys
1 5

<210> 3
<211> 9
<212> PRT
<213> G. demissa

<220>
<221> VARIANT
<222> (2)
<223> Position 2 is a Thr or Ala residue.

<220>
<221> VARIANT
<222> (4)..(6)
<223> Position 4 is a Dopamine or Tyr residue. Position 5 is a Ser, Val, Asp, or Leu residue. Position 6 is an Ala, Pro, Hydroxyproline, or Leu residue.

<220>
<221> VARIANT
<222> (8)
<223> Position 8 is a Dopamine or Tyr residue.

<400> 3
Gln Xaa Gly Xaa Xaa Xaa Gly Xaa Lys
1 5

<210> 4
<211> 7
<212> PRT
<213> Aulecomya ater

<220>
<221> VARIANT
<222> (3)
<223> Position 3 is a Tyr or Dopamine residue.

<220>
<221> VARIANT
<222> (6)..(7)
<223> Position 6 is a Val, Leu, or Ile residue. Position 7 is a Lys, or Hydroxylysine residue.

<400> 4
Ala Gly Xaa Gly Gly Xaa Xaa
1 5

<210> 5
<211> 10
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Designed peptide to act as a bioadhesive after hydroxlation.

<400> 5
Ala Lys Pro Ser Tyr Pro Pro Thr Tyr Lys
1 5 10

<210> 6
<211> 10
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Designed sequence for producing an artificial bioadhesive protein.

<220>
<221> VARIANT
<222> (3)
<223> Position 3 is a 3-hydroxyproline, 4-hydroxyproline, 3,4-dihydroxyproline, 4-fluoroproline, 3-fluoroproline, or 3,4-epoxyproline.

<220>
<221> VARIANT
<222> (6)..(7)
<223> Positions 6 and 7 are a 3-hydroxyproline, 4-hydroxyproline, 3,4-dihydroxyproline, 4-fluoroproline, 3-fluoroproline, or 3,4-epoxyproline.

<400> 6
Ala Lys Xaa Ser Tyr Xaa Xaa Thr Tyr Lys
1 5 10

<210> 7
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Designed sequence for

producing an artificial bioadhesive protein.

<220>
<221> CDS
<222> (1)..(30)

<400> 7
gct aaa ccg tcc tac cca ccg acc tac aag 30
Ala Lys Pro Ser Tyr Pro Pro Thr Tyr Lys
1 5 10

<210> 8
<211> 351
<212> DNA
<213> Artificial sequence

<220>
<223> Description of Artificial Sequence: Designed sequence for producing an artificial bioadhesive protein.

<220>
<221> CDS
<222> (19)..(318)

<400> 8
ccggccggat ccatggct gct aaa ccg tcc tac cca ccg acc tac aag gct 51
Ala Lys Pro Ser Tyr Pro Pro Thr Tyr Lys Ala
1 5 10

aaa ccg tcc tac cca ccg acc tac aag gct aaa ccg tcc tac cca ccg 99
Lys Pro Ser Tyr Pro Pro Thr Tyr Lys Ala Lys Pro Ser Tyr Pro Pro
15 20 25

acc tac aag gct aaa ccg tcc tac cca ccg acc tac aag gct aaa ccg 147
Thr Tyr Lys Ala Lys Pro Ser Tyr Pro Pro Thr Tyr Lys Ala Lys Pro
30 35 40

tcc tac cca ccg acc tac aag gct aaa ccg tcc tac cca ccg acc tac 195
Ser Tyr Pro Pro Thr Tyr Lys Ala Lys Pro Ser Tyr Pro Pro Thr Tyr
45 50 55

aag gct aaa ccg tcc tac cca ccg acc tac aag gct aaa ccg tcc tac 243
Lys Ala Lys Pro Ser Tyr Pro Pro Thr Tyr Lys Ala Lys Pro Ser Tyr
60 65 70 75

cca ccg acc tac aag gct aaa ccg tcc tac cca ccg acc tac aag gct 291
Pro Pro Thr Tyr Lys Ala Lys Pro Ser Tyr Pro Pro Thr Tyr Lys Ala
80 85 90

aaa ccg tcc tac cca ccg acc tac aag gcttaatgag gatccgaatt 338
Lys Pro Ser Tyr Pro Pro Thr Tyr Lys
95 100

caagcttcgg gcc 351

<210> 9
<211> 10
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Designed sequence for producing an artificial bioadhesive protein.

<220>
<221> HYDROXYLATION
<222> (3)
<223> Position 3 is hydroxyproline.

<220>
<221> HYDROXYLATION
<222> (6)..(7)
<223> Positions 6 and 7 are hydroxyproline.

<400> 9
Ala Lys Xaa Ser Tyr Xaa Xaa Thr Tyr Lys
1 5 10

<210> 10
<211> 10
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Designed sequenced for producing an artificial bioadhesive protein.

<220>
<221> HYDROXYLATION
<222> (3)
<223> Position 3 is a hydroxyproline residue.

<220>
<221> HYDROXYLATION, OTHER RESIDUE
<222> (5)..(7)
<223> Position 5 is a Dopamine residue. Positions 6 and 7 are hydroxyproline residues.

<220>
<221> OTHER RESIDUE
<222> (9)
<223> Position 9 is a dopamine residue.

<400> 10
Ala Lys Xaa Ser Xaa Xaa Xaa Thr Xaa Lys
1 5 10

<210> 11
<211> 219
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Designed sequenced for producing an artificial bioadhesive protein.

<220>
<221> CDS
<222> (30)..(179)

<400> 11
gggaattcac catggcgagc tcagctaaa gct aaa ccg tcc tac cca ccg acc 53
Ala Lys Pro Ser Tyr Pro Pro Thr
1 5

tac aag gct aaa ccg tcc tac cca ccg acc tac aag gct aaa ccg tcc 101
Tyr Lys Ala Lys Pro Ser Tyr Pro Pro Thr Tyr Lys Ala Lys Pro Ser
10 15 20

tac cca ccg acc tac aag gct aaa ccg tcc tac cca ccg acc tac aag 149
Tyr Pro Pro Thr Tyr Lys Ala Lys Pro Ser Tyr Pro Pro Thr Tyr Lys
25 30 35 40

gct aaa ccg tcc tac cca ccg acc tac aag gctaaatacc cctcaggtca 199
Ala Lys Pro Ser Tyr Pro Pro Thr Tyr Lys
45 50

agcttggatc ctctagaggg 219

<210> 12
<211> 168
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Designed sequence for

producing an artificial bioadhesive protein.

<220>
<221> CDS
<222> (1)..(168)
<400> 12

```
tca gct aaa gct aaa ccg tcc tac cca ccg acc tac aag gct aaa ccg     48
Ser Ala Lys Ala Lys Pro Ser Tyr Pro Pro Thr Tyr Lys Ala Lys Pro
  1               5                  10                  15

tcc tac cca ccg acc tac aag gct aaa ccg tcc tac cca ccg acc tac     96
Ser Tyr Pro Pro Thr Tyr Lys Ala Lys Pro Ser Tyr Pro Pro Thr Tyr
            20                  25                  30

aag gct aaa ccg tcc tac cca ccg acc tac aag gct aaa ccg tcc tac    144
Lys Ala Lys Pro Ser Tyr Pro Pro Thr Tyr Lys Ala Lys Pro Ser Tyr
        35                  40                  45
cca ccg acc tac aag aaa tac ccc                                   168
Pro Pro Thr Tyr Lys Lys Tyr Pro
    50                  55
```

<210> 13
<211> 336
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Designed sequence for producing an artificial bioadhesive protein.

<220>
<221> CDS
<222> (1)..(336)

<400> 13

```
tca gct aaa gct aaa ccg tcc tac cca ccg acc tac aag gct aaa ccg     48
Ser Ala Lys Ala Lys Pro Ser Tyr Pro Pro Thr Tyr Lys Ala Lys Pro
 1               5                  10                  15

tcc tac cca ccg acc tac aag gct aaa ccg tcc tac cca ccg acc tac     96
Ser Tyr Pro Pro Thr Tyr Lys Ala Lys Pro Ser Tyr Pro Pro Thr Tyr
            20                  25                  30

aag gct aaa ccg tcc tac cca ccg acc tac aag gct aaa ccg tcc tac    144
Lys Ala Lys Pro Ser Tyr Pro Pro Thr Tyr Lys Ala Lys Pro Ser Tyr
        35                  40                  45

cca ccg acc tac aag aaa tac ccc tca gct aaa gct aaa ccg tcc tac    192
Pro Pro Thr Tyr Lys Lys Tyr Pro Ser Ala Lys Ala Lys Pro Ser Tyr
```

```
     50                  55                  60

cca ccg acc tac aag gct aaa ccg tcc tac cca ccg acc tac aag gct    240
Pro Pro Thr Tyr Lys Ala Lys Pro Ser Tyr Pro Pro Thr Tyr Lys Ala
 65                  70                  75                  80

aaa ccg tcc tac cca ccg acc tac aag gct aaa ccg tcc tac cca ccg    288
Lys Pro Ser Tyr Pro Pro Thr Tyr Lys Ala Lys Pro Ser Tyr Pro Pro
                 85                  90                  95

acc tac aag gct aaa ccg tcc tac cca ccg acc tac aag aaa tac ccc    336
Thr Tyr Lys Ala Lys Pro Ser Tyr Pro Pro Thr Tyr Lys Lys Tyr Pro
            100                 105                 110
```

<210> 14
<211> 6
<212> PRT

<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Designed sequence for producing an artificial bioadhesive protein.

<220>
<221> VARIANT
<222> (3)
<223> Position 3 is a proline or hydroxyproline residue.

<400> 14

```
Lys Tyr Xaa Ser Ala Lys
  1               5
```

<210> 15
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Designed sequence for producing an artificial bioadhesive protein.

<220>
<221> CDS
<222> (1)..(18)

<400> 15

```
aaa tac ccc tca gct aaa                                                  18
Lys Tyr Pro Ser Ala Lys
  1               5
```

INTERNATIONAL SEARCH REPORT

Int tional Application No
PCT/US 99/20463

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 C12N15/12 C07K14/435

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 C12N C07K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X Y	WO 88 07076 A (GENEX CORP) 22 September 1988 (1988-09-22) page 5, line 20 -page 6, line 2; claims 1-33; example 2	1-20 25-28
A	WAITE J H: "EVIDENCE FOR A REPEATING 3,4-DIHYDROXYPHENYLALANINE-AND HYDROXYPROLINE-CONTAINING DECAPEPTIDE IN THE ADHESIVE PROTEIN OF THE MUSSEL, MYTILUS EDULIS L" JOURNAL OF BIOLOGICAL CHEMISTRY. (MICROFILMS),US,AMERICAN SOCIETY OF BIOLOGICAL CHEMISTS, BALTIMORE, MD, vol. 258, no. 5, page 2911-2915 XP000651279	1-20
Y		25-28

-/—

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
14 January 2000	28/01/2000

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL – 2280 HV Rijswijk Tel. (+31–70) 340–2040, Tx. 31 651 epo nl, Fax (+31–70) 340–3016	Nauche, S

2

Form PCT/ISA/210 (second sheet) (July 1992)

International Application No
PCT/US 99/20463

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 88 03953 A (GENEX CORP) 2 June 1988 (1988-06-02) claims 1-75	1-20
A	INOUE K ET AL: "Mussel adhesive plaque protein gene is a novel member of epidermal growth factor-like gene family." J BIOL CHEM, MAR 24 1995, 270 (12) P6698-701, XP002127537 UNITED STATES	1-20
Y	the whole document	25-28
A	US 5 410 023 A (BURZIO LUIS O) 25 April 1995 (1995-04-25)	

INTERNATIONAL SEARCH REPORT

Information on patent family members

Ir xtional Application No
PCT/US 99/20463

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 8807076 A	22-09-1988	AU 1572588 A	10-10-1988
		EP 0304486 A	01-03-1989
		JP 2500082 T	18-01-1990
		US 5242808 A	07-09-1993
WO 8803953 A	02-06-1988	AU 1180488 A	16-06-1988
		DK 414688 A	22-07-1988
		EP 0332660 A	20-09-1989
		US 5202236 A	13-04-1993
		US 5049504 A	17-09-1991
		US 5202256 A	13-04-1993
		US 5242808 A	07-09-1993
		US 5149657 A	22-09-1992
US 5410023 A	25-04-1995	NONE	